Probabilistic Models of Grammar Acquisition

Probabilistic Models of Grammar Acquisition

Mark Steedman

University of Edinburgh
steedman@inf.ed.ac.uk

The most convincing models of human grammar acquisition to date are supervised, in the sense that they learn from pairs of strings and meaning representations (Siskind, 1996; Villavicencio, 2002; Villavicencio, 2011; Buttery, 2004; Buttery, 2006; Kwiatkowski et al., 2012). Although the principles by which such models learn are quite general, the datasets they have been applied to have unavoidably been somewhat target-language-specific, and are also limited to discourse-external world-state-related content, contrary to the observations of (Tomasello, 2001) concerning the central role of common ground and grounding in interpersonal interaction.

I'll review the state of the art in the light of these limitations on the datasets, and try to make some suggestions about how we might obtain more realistic and challenging artificial and natural datasets using both automatic and human labeling methods.

References

Paula Buttery. 2004. A quantitative evaluation of naturalistic models of language acquisition: the efficiency of the Triggering Learning Algorithm compared to a Categorial Grammar Learner. In *Proceedings of the First Workshop on Psycho-computational Models of Human Language Acquisition*, pages 1–6.

Paula J. Buttery. 2006. Computational models for first language acquisition. Technical Report UCAM-CL-TR-675, University of Cambridge, Computer Laboratory.

Tom Kwiatkowski, Sharon Goldwater, Luke Zettelmoyer, and Mark Steedman. 2012. A probabilistic model of syntactic and semantic acquisition from child-directed utterances and their meanings. In *Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics*.

Jeffery Mark Siskind. 1996. A computational study of cross-situational techniques for learning word-to-meaning mappings. *Cognition*, 61:39–91.

Michael Tomasello. 2001. Perceiving intentions and learning words in the second year of life. In Melissa Bowerman and Stephen Levinson, editors, *Language Acquisition and Conceptual Development*, pages 132–158. Cambridge University Press, Cambridge.

Aline Villavicencio. 2002. The acquisition of a unification-based generalised categorial grammar. Technical Report UCAM-CL-TR-533, University of Cambridge,Computer Laboratory.

Aline Villavicencio. 2011. Language acquisition with a unification-based grammar. In R. Borsley e K. Borjars, editor, *Non-transformational Syntax*. Blackwell.

Proceedings of the EACL 2012 Workshop on Computational Models of Language Acquisition and Loss, page 19,
Avignon, France, April 24 2012.